Chapter III

Regulation of *Spblimp1/krox1a*: an alternatively transcribed isoform expressed in the mid and hindgut of the sea urchin gastrula.

Carolina B. Livi, and Eric H. Davidson

(In Press)

ABSTRACT

The sea urchin regulatory gene *Spblimp1/krox* produces alternatively transcribed and spliced isoforms, *1a* and *1b*, which have different temporal and spatial patterns of expression. Here we describe a *cis*-regulatory module that controls the expression of the *1a* splice form in the midgut and hindgut at the beginning of gastrulation. Conserved sequence patches revealed by a comparison of the *blimp1/krox* locus in *Strongylocentrotus purpuratus* and *Lytechinus variegatus* genomes were tested by gene transfer, in association with GFP or CAT reporter genes. An expression construct containing a conserved sequence patch immediately 5' of exon 1a included the transcription initiation site for *blimp1/krox1a.* This construct displays specific mid and hindgut expression, indicating that these are the locations of endogenous *blimp1/krox1a* transcription during the gastrula stage. Its sequence contains binding sites for Brn1/2/4, Otx, and Blimp1/Krox itself, as predicted in a prior regulatory network analysis.

Key Words: *Strongylocentrotus purpuratus*, *Lytechinus variegatus*, *blimp1/krox*, *Spblimp1/krox1a*, *cis*-regulation, gut specific regulatory module

1. Results and discussion

The endomesoderm of the sea urchin embryo is derived from the vegetal plate, which after the 60-cell stage consists of three different lineage compartments. Viewed from the vegetal pole, these are arranged concentrically: at the pole is the skeletogenic micromere lineage, surrounding this is the veg_2 lineage that gives rise to embryonic mesoderm and endoderm territories, and outside of $veg_2$is the veg_1 lineage, at the junction with the animal half ectoderm lineages (Davidson et al., 1998). The embryonic gut or archenteron is formed by cells from the veg_2 and the inner part of the veg_1compartments (Logan and McClay, 1997; Ransick and Davidson, 1998). Following gastrular invagination, the gut is subdivided into hindgut, midgut, and foregut, which express distinct sets of regulatory genes.

The *blimp1/krox* gene produces at least two isoforms, *blimp1/krox1a*, and *blimp1/krox1b.* Both are expressed in the endomesoderm territories. The *blimp1/krox1b* transcript appears in the vegetal plate during cleavage stages (Livi and Davidson, 2006). (1996). The *blimp1/krox* gene is expressed in blastopore and hindgut early in gastrulation, when both isoforms are present, and its expression soon extends into the midgut (Livi and Davidson, 2006). The late or *1a* isoform was first described by Wang et al. Here we focus on the regulation of this isoform, which begins to be expressed only after 30 hours postfertilization (hpf), coincident with the onset of gastrulation.

An extensive endomesoderm gene regulatory network (EM-GRN) indicating the functional relationships between regulatory genes within the vegetal plate up to 30 hours of development has been published (Davidson et al., 2002a; b; Oliveri and Davidson, 2004; Levine and Davidson, 2005). The network architecture of the postgastrular archenteron is starting to take shape as well (Yuh et al., 2005). *Spblimp1/krox* appears to be part of this network, and is likely to be involved in the separation of veg_1- derived endoderm from veg_1- derived ectoderm, since it is expressed only in the prospective endoderm portions.

1.1. Genomic sequence surrounding the blimp1/krox gene

The *blimp1/krox* gene extends over 42 kb. An *Spblimp1/krox* BAC (163O19) described earlier (Livi and Davidson, 2006) is 94 kb in length, and contains 43 kb of upstream sequence, all 7 *blimp1/krox* exons and their 6 introns, and 11 kb of downstream sequence. The *Lvblimp1/krox* BAC (60B16) used for comparison in the following is 54 kb in length, and contains 10 kb of upstream sequence, 6 identifiable *blimp1/krox* exons and 5 introns, and 10 kb of downstream sequence. The alignment of these BACs is shown in Fig. 1A. Exon 6 cannot be identified in the *Lv* BAC. Positions 36k to 91k on the *Spblimp1/krox* BAC correspond to the region included in the *Lvblimp1/krox* BAC.

The alternatively transcribed isoforms are conserved between *S. purpuratus* and *L. variegatus*, including the 5' UTR sequence encoded in exon *1a*. However, there are no

expression data for the late isoform in *L. variegatus*, so we cannot confirm that it is expressed in a similar manner in the two species.

1.1.2. Identification by "Family Relations" analysis of conserved genomic sequence patches in the vicinity of the Spblimp1/krox *gene*

Comparison of *S. purpuratus* and *L. variegatus* DNA sequence in the vicinity of given genes has successfully revealed the location of *cis*-regulatory modules controlling specific aspects of expression of these genes (e.g., Yuh et al., 2002; Revilla-i-Domingo et al., 2004; Minokawa et al., 2005). These two sea urchin species diverged approximately 50 mya (Smith, 1988; Lee, 2003). Using the "Family Relations" software of Brown et al. (2002, 2005), analyses of the overlapping *Spblimp1/krox* BAC regions produced a set of conserved sequence patches that may be considered to be putative regulatory elements. Their location is shown in Fig. 1B. Conserved sequence patches in the large *1a-1b* intron, 3' of position 50kb on the *Spblimp1/krox* BAC, control expression of the early form, as will be described elsewhere (J. Smith, C. Livi, and E. Davidson, unpublished data). The longest and most conserved patch (761 bp) is located 5' of exon 1a, at 43kb on the *Spblimp1/krox* BAC, and at 8 kb on the *Lvblimp1/krox* BAC (Fig.1C).

1.2. A DNA fragment which accurately generates the late midgut and hindgut pattern of blimp1/krox *gene expression*

An 850 bp construct including the conserved non-coding genomic sequence immediately 5' of exon *1a* plus the 5' UTR of exon *1a*, and a CAT reporter gene (or a GFP reporter gene), was created by fusion-based PCR (Construct 43, Fig. 1C; see Experimental procedures for details). The linear PCR product was injected into sea urchin zygotes. No significant expression in blastula stage embryos was seen (data not shown). We have previously shown using QPCR that *Spblimp1/krox1a* starts being expressed after 30 hpf (Livi and Davidson, 2006).

1.2.1. Spatial expression of 43-CAT and 43-GFP constructs

The endogenous *Spblimp1/krox* gene is expressed in the midgut, hindgut, and blastoporal region of the gastrula stage embryo (Livi and Davidson, 2006; Fig. 2J-L). The spatial pattern of expression driven by Construct 43 can be considered to represent the *1a* late form because it derives from sequences immediately adjacent to the *1a* exon with which the late transcript begins. To determine this pattern of expression, whole-mount in situ hybridization (WMISH) was carried out at mid-gastrula stage using antisense CAT or GFP probes. This has the advantage over GFP fluorescence that the mRNA is less stable than is GFP protein, so the results do not sum over time. However, GFP fluorescence can be usefully monitored in live embryos for initial screening of each experimental batch, and when desired, the embryos fixed for WMISH. Fig. 2A-C show WMISH displays of transgenic embryos bearing construct 43-GFP; Fig. 2D-I show embryos bearing construct 43-CAT. Transgene incorporation is mosaic in sea urchins (Hough-Evans et al., 1988; Livant et al., 1991); thus, to interpret the complete pattern, it is

necessary to visualize multiple integration events by looking at several embryos. As Table 1 shows, 95% of gastrula stage embryos displayed robust expression in the gut. Therefore, the endogenous *blimp1/krox* transcript expression at these stages in the midgut and hindgut is probably of *1a* origin. No differences were observed in analyzing embryos injected with the two different module 43 reporter constructs.

1.2.2. Annotation of 900 bp sequence immediately 5' of exon 1a, including putative transcription factor binding sites

The region including the conserved patch is shown as underlined text in Figure 3. Yellow highlights correspond to the location of primers used to amplify or check the fusion construct (see Experimental Procedures for primer sequences and other details). Exon 1a is highlighted in gray, and is comprised exclusively of 5' UTR. The ATG for this isoform is located in exon 2 (Livi and Davidson, 2006). The locations of some putative transcription factor binding sites are also indicated in Fig. 3, viz. sites for Brn1/2/4, Otx, and Blimp1/Krox itself (see legend). One of the Otx sites lies just outside the conserved region, but all others are within it. These sites are not all conserved in the same positions in *Lv* as in *Sp*, however.

A 709 bp conserved patch is found 3' of exon 1a (Fig. 1C), which was not included in construct 43. Given that construct 43 conferred proper temporal and spatial expression, and that it contains target sites for the inputs predicted by the EM-GRN, we

did not pursue this region further. Possibly, it has regulatory function in other stages of development.

1.2.3. cis-*Regulatory analysis of expression driven by module 43 and the network model for endomesoderm specification*

The *blimp1/krox* gene encodes a transcriptional regulator centrally positioned in the EM-GRN, where it performs a required early role in the specification of the endomesoderm and subsequently the endoderm of the embryo (Davidson et al., 2002a; b; Livi and Davidson, 2006). Several of its direct downstream targets have been identified and confirmed by *cis*-regulatory analysis (Yuh et al., 2004; Minokawa et al., 2005). These relationships pertain to the early *1b* isoform. We have now generated predictions for the predicted inputs into the *1a* late form as well, *viz*. the factors encoded by the *otx*, and *brn1/2/4* genes and the *blimp1/krox* gene product itself (Yuh et al., 2005; C.-H. Yuh and E. Davidson, unpublished data). These predictions were based on perturbation analyses, as well as time and place of expression of the respective genes. To further our understanding of the control system that determines expression of the late isoform, we have here identified the *cis*-regulatory element responsible for driving its midgut- and hindgut-specific expression. Fig. 3 shows that the target sites that could mediate all the predicted inputs into this control system are indeed found within the conserved sequence patch that confers the correct spatial and temporal expression pattern. One point of interest is that though the early isoform is also subject to autoregulation by the *blimp1/krox* gene

product, the negative feedback loop in which it participates (Livi and Davidson, 2006) does not appear to apply to the late isoform.

2. Experimental procedures

2.1. Isolation and analysis of BAC clones containing Spblimp1/krox *and* Lvblimp1/krox *genes*

Spblimp1/krox (clone 163O19) and *Lvblimp1/krox* (clone 60B16) BACs were obtained by hybridization of a *Spblimp1/krox* cDNA fragment to arrayed *S. purpuratus* and *Lytechinus variegatus* (*L. variegatus*) BAC libraries using a probe generated by PCR amplification of the *S. purpuratus* gene (left primer: 5'-ACAACAGCTGCAGGGACATT-3'; right primer: 5'-ACATGGTCCGACACACTGAT-3'). The fragment was subcloned into pGEMTeasy (Invitrogen, Carlsbad, CA). *Spblimp1/krox* and *Lvblimp1/krox* BACs were sequenced by the DOE Joint Genome Institute (Genebank accession nos. **AC131508**; **AC131502;** Cameron et al., 2000; 2004).

2.2. Comparison of the genomic sequence around the blimp1/krox *genes of* S. purpuratus *and* L. variegatus

The BAC sequences were annotated using the Sea Urchin Genome Annotation Resource (SUGAR), as well as Family Relations (Brown et al., 2002; 2005). The exons of the *blimp1/krox* gene in *S. purpuratus* and *L. variegatus* BAC clones were identified

using the sequence of *S. purpuratus* and *L. variegatus* cDNA clones (Genebank accession nos. **DQ225099; DQ177152**). The sequences were annotated using the SUGAR software package (Brown et al., 2002). This software was used to identify coding sequences of genes neighboring *blimp1/krox* in the BAC clones. The Family Relations software package (Brown et al., 2002; 2005) was used to compare the genomic BAC sequences of *S. purpuratus* and *L. variegatus*. Window sizes used in the comparison ranged from 10 to 100 bp, and the similarity values ranged from 70 to 100%. The pairwise view of the software was used to identify conserved regions. The Dot Plot view was used in some cases to identify the boundaries of the conserved regions found.

2.3. Binding site searches

Putative binding site sequences were obtained from different sources and searched using the Family Relation Software package (Brown et al., 2002; 2005). The consensus binding sites searched corresponded to the following:

Blimp1: 5' - G(A/G)AA(C/G)(GT)GAAA -3' or 5' - G(A/G)AA(C/G)(GT)AAA - 3' (Gupta et al., 2001; Marecki and Fenton, 2002, Yuh et al., 2004);

Otx: 5' - TAATC(C/T) - 3' (Yuh et al., 2005);

Brn1/2/4: 5' - GCAT(N_{0-3})TAA(A/T) - 3' (Yuh et al., 2005).

The Brn1/2/4 binding site in the *endo16* promoter (endo16 site=5' – GCAT(GAA)TAAA - 3') differs from the previously reported canonical site (canonical=5' - GCAT(N_{0-3})TAAT) - 3') in the same way as the site found in the

blimp1/krox construct 43 (blimp1/krox site= 5' – GCAT()TAAA- 3'). However, it lacks the middle nucleotides present in the *endo16* Brn1/2/4 site. We changed the canonical site to reflect the sea urchin Brn1/2/4 sites that have been found to 5' - GCAT(N_{0-3})TAA(A/T).

2.4. Preparation of reporter constructs

All sequences were amplified using the "Expand High Fidelity PCR System" (Roche, Indianapolis, IN). Promoter sequences were amplified from BAC 163O19 (**AC131508**). GFP sequence was amplified from the pGreenL3 vector (Arnone et al., 2004). CAT sequences were amplified from the CE CAT vector (Yuh et al., 2002). Expression constructs were prepared by fusion PCR as previously described (Hobert, 2002). The sequence of primers used is listed below. The first set of primers is used to amplify the individual fragments. The internal fusion primers are used in the presence of the two fragments to create a contiguous sequence. PCR product itself was injected into the embryos after purification using a Zymoclean Gel DNA Recovery Kit (Zymo Research, Orange, CA).

Fragments 43:

First43.4k_For: 5'-ATCTCCCAACGACAAGATGG-3'

BP1a_Rev_GFP: 5'-

GCTAGCACGCGTAAGAGCTCGGTAACCTCGTTGAAGCGTAG-3'

BP1a_Rev_CAT: 5'-

CTCCATTTTAGCTTCCTTAGCTCCTGAACCTCGTTGAAGCGTAG-3'

CAT Fragment:

CAT_For: 5'-TCAGGAGCTAAGGAAGCTAAAATGGAG-3'

firstCAT_Rev: 5'-CTATAGGGCGAATTGGAGCTCCAC-3'

GFP Fragment:

GFP_For: 5'-TACCGAGCTCTTACGCGTGCTAGC-3'

firstGFP_Rev: 5'-CACCTGTCCTACGAGTTGCATG-3'

Fragment 43 CAT Fusion (internal primers):

Nested43.4k_For: 5'-CTAGCTTCGTGTTGCAAATTG-3'

nestedDD-CAT B- 5'-TGAGTTTGGACAAACCACAAC-3'

Fragment 43 GFP Fusion (internal primers):

Nested43.4k_For: 5'-CTAGCTTCGTGTTGCAAATTG-3'

NestedGFP_REV: 5'-CTGACTGGGTTGAAGGCTCTC-3'

2.5. Microinjection of reporter constructs, whole mount in situ hybridization and embryo imaging

Gamete collection and microinjection were performed as previously described (Livi and Davidson, 2006).

Whole-mount in situ hybridization and embryo image acquisition were performed as previously described by Livi and Davidson (2006). Briefly, digoxigenin-labeled RNA probes corresponded to the antisense as well as the sense direction. No staining was observed using the sense probes (data not shown). *Spblimp1/krox* antisense probe is complementary to sequence in exon 2 present in both isoforms (Livi and Davidson, 2006).

2.6. Accession numbers

Genebank accession nos. are **AC131508**; **AC131502; DQ225099; DQ177152.**

Acknowledgments

We thank Takuya Minokawa for assistance with fusion PCR, and for sequences for CAT fusion primers and templates, and Titus Brown for assistance with the Family Relations analysis. This work was supported by a grant from the National Institute of Child Health and Human Development.

References

Arnone, M. I., Dmochowski, I. J., Gache, C., 2004. Using reporter genes to study cis-regulatory elements. Methods Cell Biol.74, 621-652.

Brown, C.T., Rust, A.G., Clarke, P.J., Pan, Z., Schilstra, M.J., De Buysscher, T., Griffin, G., Wold, B.J., Cameron, R.A., Davidson, E.H., Bolouri, H., 2002. New computational approaches for analysis of cis-regulatory networks. Dev Biol. 246, 86-102.

Brown, C.T., Xie, Y., Davidson, E.H., Cameron, R.A., 2005. Paircomp, FamilyRelationsII and Cartwheel: tools for interspecific sequence comparison. BMC Bioinformatics. 24, 70-76.

Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R., Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., Wray, G.A., Ettensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H., Hood, L., 2000. A sea urchin genome project: sequence scan, virtual map, and additional resources. Proc. Natl. Acad. Sci. U S A. 97, 9514-9518.

Cameron, R.A., Oliveri, P., Wyllie, J., Davidson, E.H., 2004. *cis*-Regulatory activity of randomly chosen genomic fragments from the sea urchin. Gene Exp. Patterns. 4, 205-213.

Davidson E.H., Cameron RA, Ransick A., 1998. Specification of cell fate in the sea urchin embryo: summary and some proposed mechanisms. Development. 125, 3269-3290.

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Schilstra, M.J., Clarke, P.J., Rust, A.G., Pan, Z., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002. A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. Dev. Biol. 246, 162-190.

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A.G., Pan, Z., Schilstra, M.J., Clarke, P.J., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002. A genomic regulatory network for development. Science. 295, 1669-1678.

Gupta, S., Anthony, A., Pernis, A.B., 2001. Stage-specific modulation of IFN-regulatory factor 4 function by Krüppel-type zinc finger proteins. J. Immunol. 166, 6104-6111.

Hobert, O., 2002. PCR fusion-based approach to create reporter gene constructs for expression analysis in transgenic C. elegans. Biotechniques. 32, 728-730.

Hough-Evans, B.R., Britten, R.J., Davidson, E.H., 1988. Mosaic incorporation and regulated expression of an endogenous gene in the sea urchin embryo. Dev. Biol. 129, 198-208.

Lee, Y.H., 2003. Molecular phylogenies and divergence times of sea urchin species of Strongylocentrotidae, Echinoida. Mol. Biol. Evol. 20, 1211-1221.

Levine, M., Davidson, E.H., 2005. Gene regulatory networks for development. Proc. Natl. Acad. Sci. USA 102, 4936-4942.

Livant, D.L., Hough-Evans, B.R., Moore, J.G., Britten, R.J., Davidson, E.H., 1991. Differential stability of expression of similarly specified endogenous and exogenous genes in the sea urchin embryo. Development. 113, 385-398.

Livi, C.B., Davidson, E.H., 2006. Expression and function of *blimp1/krox*, an alternatively transcribed regulatory gene of the sea urchin endomesoderm network. Dev. Biol. In press.

Logan, C.Y., McClay, D.R., 1997. The allocation of early blastomeres to the ectoderm and endoderm is variable in the sea urchin embryo. Development. 124, 2213-2223.

Marecki, S., Fenton, M.J., 2002. The role of IRF-4 in transcriptional regulation. J Interferon Cytokine Res. 22, 121-133.

Minokawa, T., Wikramanayake, A.H., Davidson, E.H., 2005. *cis*-Regulatory inputs of the *wnt8* gene in the sea urchin endomesoderm network. Dev. Biol. 288, 545-558.

Oliveri, P., Davidson, E.H., 2004. Gene regulatory network controlling embryonic specification in the sea urchin. Curr. Opin. Genet. Dev. 14, 351-360.

Ransick, A., Davidson, E.H., 1998. Late specification of veg_1 lineages to endodermal fate in the sea urchin embryo. Dev. Biol. 195, 38-48.

Revilla-i-Domingo, R., Minokawa, T., Davidson, E.H., 2004. R11: a cis-regulatory node of the sea urchin embryo gene network that controls early expression of SpDelta in micromeres. Dev. Biol. 274, 438–451.

Smith, A.B., 1988. Phylogenetic relationship, divergence times, and rates of molecular evolution for camarodont sea urchins. Mol. Biol Evol. 5, 345-365.

Wang, W., Wikramanayake, A.H., Gonzalez-Rimbau, M., Vlahou, A., Flytzanis, C.N., Klein, W.H., 1996. Very early and transient vegetal-plate expression of SpKrox1, a Kruppel/Krox gene from Stronglyocentrotus purpuratus. Mech. Dev. 60, 185-195.

Yuh, C.H., Brown, C.T., Livi, C.B., Rowen, L., Clarke, P.J., Davidson, E.H., 2002. Patchy interspecific sequence similarities efficiently identify positive cis-regulatory elements in the sea urchin. Dev. Biol. 246, 148-161.

Yuh, C.H., Dorman, E.R., Howard, M.L., Davidson, E.H., 2004. An otx cis-regulatory module: a key node in the sea urchin endomesoderm gene regulatory network. Dev. Biol. 269, 536-551.

Yuh, C.H., Dorman, E.R., Davidson, E.H., 2005. Brn1/2/4, the predicted midgut regulator of the *endo16* gene of the sea urchin embryo. Dev. Biol. 281, 286-298.

Table 1

WMISH Assays of late *Spblimp1/krox1a* construct 43

Stage	Experiment (batch)	1	2	3	combined
40h	Nonstained	26	36	23	85
	Total stained embryos	63	72	45	180
	(% stained embryos)	(70%)	(67%)	(66%)	(68%)
	Gut	94%	96%	95%	95%
	(stained cells/embryos)	*4*	*4*	*4*	*4*
	Secondary mesenchyme	11%	21%	18%	16%
	(stained cells/embryo)	*2*	*1.5*	*2*	*1.9*
	Skeletogenic mesenchyme	3%	0	0	1%
	(stained cells/embryo)	*1.5*	*0*	*0*	*1.5*
	Aboral ectoderm	11%	7%	4%	5%
	(stained cells/embryo)	*2*	*2.5*	*2*	*2.1*
	Oral ectoderm	8%	3%	0	4%
	(stained cells/embryo)	*2*	*2*	*0*	*2.1*

Figure legends

Fig. 1. Comparative Family Relations II (FRII) analysis of genomic BAC sequences surrounding *Spblimp1/krox* and *Lvblimp1/krox*. (A) BAC maps of the *blimp1/krox* locus and its two transcriptional units, modified from Livi and Davidson (2006). Boxes indicate the location of exons, which are numbered. The *Spblimp1/krox* BAC, 163O19, is 94 kb long. The *Lvblimp1/krox* BAC, 60B16, is 54 kb long. (B) FRII Analysis of BAC sequences in region of overlap between the *Sp* and *Lv* BACs. A pairwise view is reproduced, using a 50 bp window at 70% similarity. Top line represents *Spblimp1/krox* BAC, while the bottom line represents *Lvblimp1/krox*. Cyan-colored boxes indicate the position of *blimp1/krox* exons. Blue-colored lines between sequences indicate patches where more than 70% of the nucleotides within the 50 base pair window are conserved. Red box indicates region shown in Fig. 1C. The structure of the spliced *Spblimp1/krox1a* cDNA is shown above, while the structure of the predicted *Lvblimp1/krox1a* cDNA is shown below the FRII analysis. (C) The same FRII Analysis as a close up of the conserved patch around exon 1a. The magenta-colored bar at the bottom indicates the location of construct 43, the subject of this paper. The sequence and annotation of fragment 43 can be seen in Fig. 3.

Figure 1

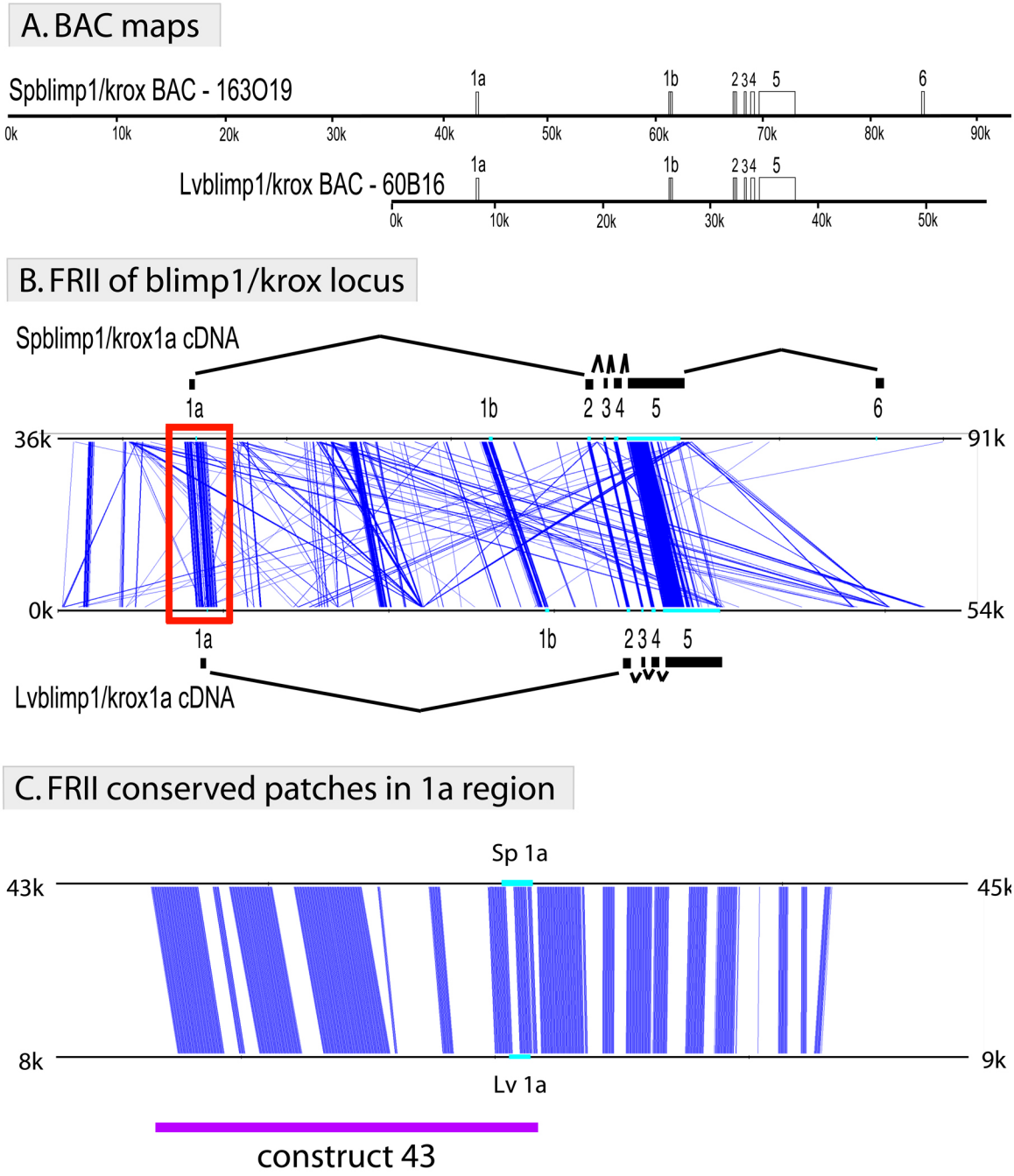
A. BAC maps
Spblimp1/krox BAC - 163O19
1a
1b
2 3 4
5
6
0k
10k
20k
30k
40k
50k
60k
70k
80k
90k
Lvblimp1/krox BAC - 60B16
1a
1b
2 3 4
5
0k
10k
20k
30k
40k
50k
B. FRII of blimp1/krox locus
Spblimp1/krox1a cDNA
1a
1b
2 3 4
5
6
36k
91k
0k
54k
1a
1b
2 3 4
5
Lvblimp1/krox1a cDNA
C. FRII conserved patches in 1a region
Sp 1a
43k
45k
8k
9k
Lv 1a
construct 43

Fig. 2. Transgene expression, and endogenous *blimp1/krox1a* expression in gastrula stage embryos. Mid-gastrula stage embryos were imaged in glycerol under Nomarsky optics. (A-C) WMISH of embryos injected with construct 43-GFP using antisense *gfp* RNA probe showing transgenic pattern of expression. (D-I) WMISH of embryos injected with construct 43-CAT using antisense *cat* RNA probe showing transgenic pattern of expression. (J-L) WMISH of uninjected embryos using antisense *blimp1/krox* RNA probe showing endogenous pattern of expression.

Figure 2

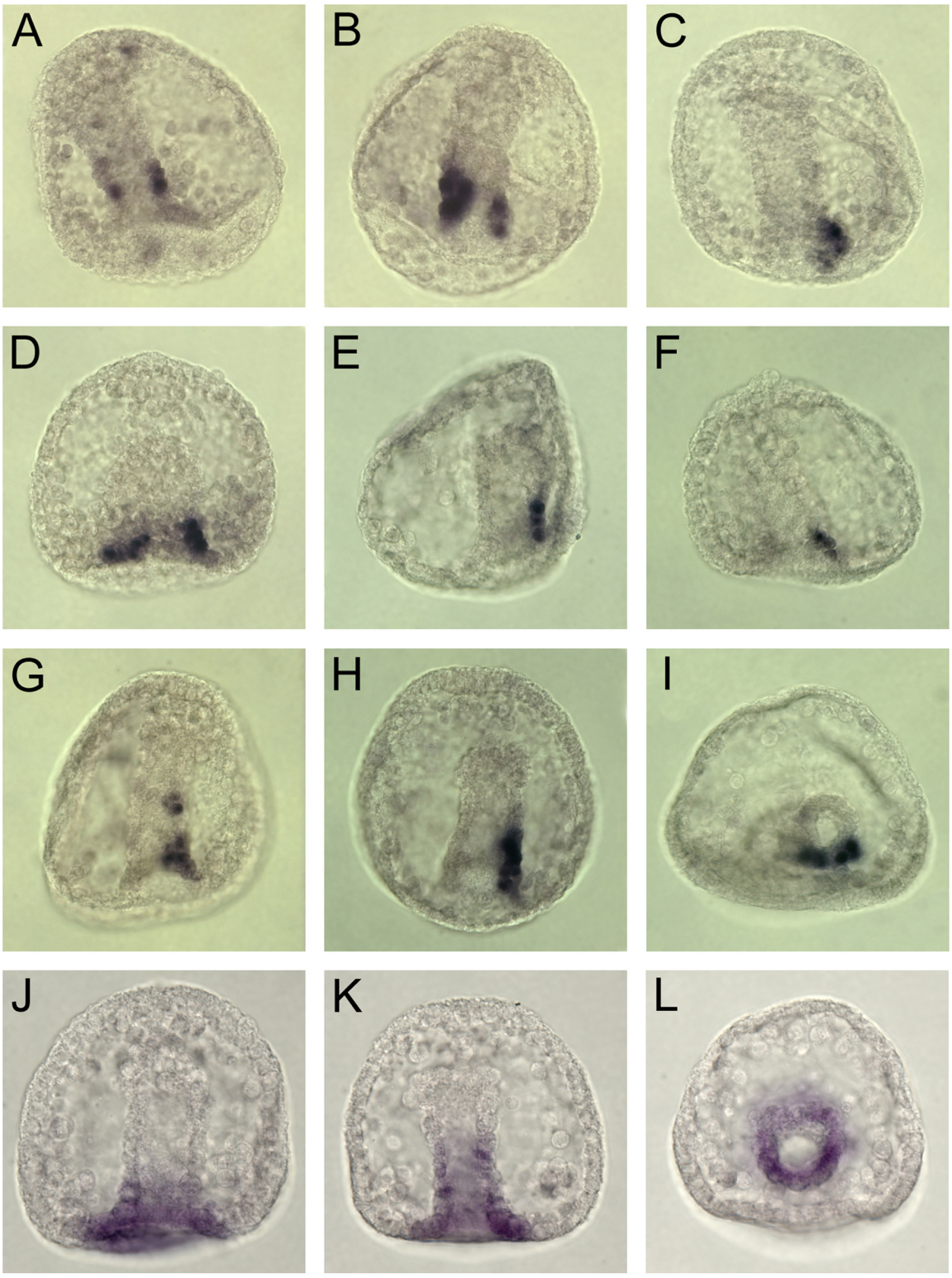

Fig. 3. Sequence of fragment 43. Underlined sequence represents the position of a conserved patch according to the Family Relations analysis. The sequence highlighted in grey corresponds to *blimp1/krox* exon 1a. The sequence highlighted in yellow indicates the location of primers utilized to amplify the sequences for subcloning or fusion PCR. Red letters correspond to the location of a putative Blimp1/Krox binding site. Magenta letters correspond to the location of a putative Brn1/2/4 binding site. Orange letters correspond to the location of putative Otx binding sites. The conserved patch extends 3' of exon 1a. This sequence is not included in construct 43, and it is not known whether or not it has regulatory activity. However, fragment 43 drives accurate spatial and temporal expression in the mid and hindgut of gastrula stage embryos.

Figure 3


TATATGAAATGAGGATTACGTAACCGATTTTTTTTGTTTTGACCTATCATTAGAGAGGTGAAAAAAATATGT
AAAAGAGGATTATGAGAAGAGCAAGAGGACGACATATGGAACAAAACTGAATAACGATAAAACGACGGTCCA
GTTTTTAATGGGTGTACGGGTGGTAACCCCCCCCCCCCCACTGTATATTCCGCCCCCCCCCCCCACACGGGG
TTTAATGTTGTGCGAAATCTCCCAACGACAAGATGGGAGGGGGGAGGGGGTTCCTAGCTTCGTGTTGCAAAT
TGTATCAGACTTTAAGTTTTAAGACCTCAAATTCTTTTCCGGGCTTTTAATCGAGGGGGGGGGGGGGCGTGG
TAAACTTTGTCGGCCAAAACCTTCGCTGTAAACGTCATTGATGGGAGAAGCTATTGCCGAAAACGATGATGA
CGACGATGATTGATGAAACCATCTTTACGTCATCAAAAATAAAGAAAAGACTAAGGTCTACGGATTTTTTTA
AAAGTAGAGTATGCGTGCCGCACAAAAACCTTCGAATGAACCGCATATTTGAGATCTGGTTTGCAGATAACT
TTCTTATCAGACCTCTCAGTAAATCTGATGCCGTGACTTCGGGCGGGTGATTACTCGAGTCGTGCTTGGACT
TTACCCCCATGAAAACTAAAAAACCAAAAGGTGAAAAACAAGCTCCAAGCTGAAGTCAGGCGGCCATCCAGA
AAGCGAAGTTGATGGTTAACCAGTCAACAAACACTTGCATTAAAACGATCAGGGTTCCCCCAGAGCCTTGTA
CCGCTATCTCTTGCGAATGTACACAAGGATAATATAAGCACGTGAACATGTATAAAAGTCAGTTTAATGACG
CTCCTTTTTCTACGTTTCTTATCTATTTATATCTAATCCAAGAAGTAACGAAAGAAATGGGAGAGAGAGGTA
GAGAGAGGGAAGGAGTGGGAGAGAGAGAGAGAGAGAGAGACGTGAAGGGTAGAAATTATCGTACTAGGCAGT
GTTAAGAGCGATTCGATCGGATACAGACGACATATTAGTAGATTTCACTTATCGGGAGAGTGCAGCGTACAG
GTGAGCGGCCGCCATTTTCTTGCACCCGCCGAAGTAAACCTATCTATCAAGTTGGAAAGCACTTGCTTGCTG
TTACCACCACAGAAAAGAGAAAAAGCGAAAAAAAGAAGTCTAAAGAAATTCCCCGATTGAAACCGCAAGCTA
TTTTTTCTACGCTTCAACGAGGTATGTCTGGTATCAGTTATTATTTTTGGAGAAAAAAATGTATTCATACAA
ACACAAATCCCTTTCTTTTTATTCACAGCGAGTCTCTTCTTTTCATTCCCCAAACCCCAGTATCTGGTTGTG
TGTGTGAGTGTGTGTGTACTTTCTGAATCTGCAATTGCTATTGCATGTGTGTGTGTGTGGGTCTGAAGTTGA
ACCGTACTTGGCCTATTGTACTGGGGGGAATAAATTGAAAGCGGTGTGCGATAATTTTTTTCATCATTTTTT
ACTTTTGCTAGTCGCAATCGCTTCCTCAAAAACAAACTTCACTTACCCGGCCGTATAGGAATCATATTACAC
AGCCGTCGCCTAAATCCTCACGACGTGGCGCTCACACGCTAAGGAGAGAAAGTTCCTCCGAAACTACAATCT
CAGGTAGATTCTAACTTCTTGGCGAGTTTTGGCGGGGTGAAAGTTTTTTTCTCTGTTGGTGAAACTTCGGGT
CGGATTTTGGCATTGGTTTGAGGCGTCGCCCTGCAGTGTATGATACCGTGTTCATTGGAAGCTAATAGGTAA
GGCCGTATTTGTGTAAGATCGTCCCCATTTTTGTTAGCATTCATATCTGTTTGTACGTGTGCGCTCAAGTTC
GGCCTATACCCCTTGCGAGAAAGTAAATGAAAATGGCGTGGTGAGAGAGAGTTCCTGAAAGTGAAC